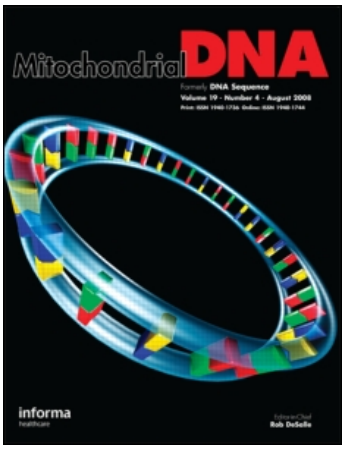

Mitochondrial DNA
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713640135

Mapping two genes in the purine metabolism pathway of soybean

J. L. Shultz [a]; J. D. Ray [a]; J. R. Smith [a]
[a] Crop Genetics and Production Research Unit, USDA-ARS, Stoneville, MS, USA

First Published:June2008

To cite this Article Shultz, J. L., Ray, J. D. and Smith, J. R.(2008)'Mapping two genes in the purine metabolism pathway of soybean',Mitochondrial DNA,19:3,264 — 269
To link to this Article: DOI: 10.1080/10425170701607522
URL: http://dx.doi.org/10.1080/10425170701607522

DNA Sequence, June 2008; 19(3): 264–269

informa
healthcare

FULL LENGTH RESEARCH PAPER

Mapping two genes in the purine metabolism pathway of soybean*

J. L. SHULTZ, J. D. RAY, & J. R. SMITH

Crop Genetics and Production Research Unit, USDA-ARS, P.O. Box 345, Stoneville, MS, USA

(Received 21 February 2007; accepted 30 July 2007)

Abstract
Mapping genes in biochemical pathways allow study of the genomic organization of pathways and geneic relationships within these pathways. Additionally, molecular markers located within the boundaries of a specific gene sequence represent important marker assisted selection resources. We report map locations of two geneic markers from the purine synthesis pathway in soybean (*Glycine max* (L. merr.)), utilizing a 90 plant F_2 population created from the cross of "DT97-4290" × "DS97-84-1". Primers were designed based on sequences from annotated soybean complimentary DNA. A polymorphic, co-dominant, sequence-characterized amplified region marker was created for hypoxanthine phosphoribosyl transferase (EC 2.4.2.8). Linkage analysis placed this gene on linkage group (LG) O. In addition, a single-nucleotide polymorphism (SNP) marker was developed for a urate oxidase gene (EC 1.7.3.3). Linkage analysis of the SNP placed the urate oxidase gene on LG I. For both genes, amplicon sequence data confirmed the identification of the respective gene. Mapping these genes represents the first step in understanding the genomic organization of the purine biochemical pathway in soybean.

Keywords: *Purine, complimentary DNA, linkage group, urate oxidase, soybean, SNP*

Introduction

A primary reason for developing molecular maps is to identify selectable markers in close proximity to genes controlling a trait of interest. Often this is necessary as the gene(s) controlling the trait are unknown. However, with the advent of high throughput sequencing, a large number of gene sequences became available. Placing genes on molecular marker maps may not only allow the identification of genes controlling a particular trait, but it may also allow a greater understanding of the mechanisms resulting in the expression of the trait, especially for complex (multi-geneic) traits (Shopinski et al. 2006). Such understanding could lead to the development of improved cultivars.

Purine biosynthesis is a central metabolic function in all cells and provides products ranging from bases for DNA and RNA to a number of essential coenzymes and signaling molecules (Smith and Atkins 2002). In legumes such as soybean, the purine pathway plays a dominant role in primary N metabolism (Atkins and Smith 2000). In soybean nodules, atmospheric N is ultimately incorporated into ureideic compounds through the purine pathway. Ureides are exported out of the nodules and provide the primary N source for the soybean plant. However, it has been demonstrated that the nitrogen fixation process may limit soybean yield and that the process is sensitive to drought (King and Purcell 2005; Ray et al. 2006). Several studies have indicated that the drought sensitivity is linked to urieide catabolism occurring in the leaves of soybean plants (Sinclair et al. 2001; Vadez and Sinclair 2001; King and Purcell 2005; Todd et al. 2006).

Mapping genes involved in purine metabolism may provide insights into understanding the limitations of N-fixation in soybean and indicate ways to overcome

Correspondence: J. D. Ray, USDA-ARS, 141 Experiment Station Road, Stoneville, MS 38776, USA. Tel: 1 662 686 3036. Fax: 1 662 686 5218. E-mail: jray@ars.usda.gov

*Mention of a trademark or proprietary product does not constitute a guarantee or warranty of the product by the US Department of Agriculture and does not imply approval or the exclusion of other products that may also be suitable.

ISSN 1042-5179 print/ISSN 1029-2365 online © 2008 Informa UK Ltd.
DOI: 10.1080/10425170701607522

the limitations. However, this is a major undertaking as 38 ESTs associated with purine metabolism have been identified in soybean (Kegg 2006). Nonetheless, as a first step in this process, we report geneic markers and map locations of two genes in the purine pathway; hypoxanthine phosphoribosyltransferase (HPT; EC 2.4.2.8) and urate oxidase (EC 1.7.3.3). HPT converts inosine-P to hypoxanthine and urate oxidase reduces urate to hydroxyisourate. Neither HPT or urate oxidase has been previously mapped with a geneic marker. However, a marker for the promoter region of the "Ur2" urate oxidase (AB002807, Soybase) has been reported (Takane et al. 1997).

Materials and methods

Mapping population and primer design

The parents, mapping population, and map framework used in this study have been detailed elsewhere (Shultz et al. 2007). Briefly, the soybean map was created from 90 F_2 plants derived from a cross between a germplasm line "DT97-4290" (Paris et al. 2006) and a breeding line "DS97-84-1" (Smith, unpublished). The F_2 plants and framework map were used to determine the genomic location of the two genes mapped in this study.

Primers were designed using a three step process. First, complimentary DNA (cDNA) sequences directly attributed to the specific gene function (via EC# and annotation) were identified on NCBI. Two sequences exhibiting polymorphisms were selected when possible. Primers were designed from these sequences using the primer 3 program (Rozen and Skaletsky 2000). Settings were an 18–27 bp oligo length, $T_m = 53–57°C$, and GC% of 20–80%. Final selection of primers was confirmed by sequencing the respective amplicons to ensure gene fidelity.

Primers for HPT were designed by comparing two soybean cDNA sequences annotated as "HPT" from NCBI. A "Williams" cDNA clone (gi4293373) and a "Harosoy" cDNA clone (gi10847241) were used to design 3 primers. Primer combinations were evaluated and ultimately, the primer pair (hypX-2 and hypX-3; Table I) detected a polymorphism in the mapping population and was used as a sequence-characterized amplified region-type marker to map HPT.

Primers for urate oxidase were designed by comparing 15 *Glycine max* EST sequences against a full length (4783 bp) cDNA clone from cv "Dare" (gi:170030). The identification and utilization of both highly conserved and polymorphic sequences led to 25 unique primers. Sequence analysis of products from combinations of primers resulted in the identification of a single-nucleotide polymorphism (SNP) suitable for use in the mapping population. The primers used for the SNP and applied to the mapping population are given in Table I.

Genotyping

For both genes, the same initial PCR protocol was employed. Components in each reaction included 1 μl 10 × PCR buffer (10 × KTLA, DNA Polymerase Technology, St Louis, MO), 0.45 μl 10 mM dNTPs (cat #U151b, Promega, Madison, WI), 0.2 μl 25 mM $MgCl_2$ (cat #c00826, Sigma Aldrich, St Louis, MO), 1.5 μl H_2O and 1.5 U Taq (KT1, DNA Pol. Tech.). A total of 5 μl target DNA (10 ng/μl) and 1 μl each of 10 mM forward and reverse primers completed the reaction. Amplification conditions consisted of an initial 94°C denaturation for 5 min followed by 35 cycles of denaturing (94°C for 60 s), annealing (57°C for 45 s) and extension (72°C for 45 s). PCR was completed by a final extension of 5 min at 72°C.

For the HPT gene, the PCR product was separated on 1.5% (g/l) agarose (cat #BP160, Fisher Scientific, Hampton, NH) prepared with ethidium bromide and electrophoresed at 4 v/cm for 3–4 h. Amplicons were documented with an Alphaimager 3000 (Alpha Innotech, San Leandro, CA).

For the urate oxidase SNP analysis, each allele specific primer extension (ASPE) SNP reaction was carried out on 5 μl of PCR product (see above) treated with 2 μl EXOSAPIT (USB cat 78200) at 37°C for 30 min, followed by deactivation for 15 min at 80°C. To each of these 7 μl samples, a total of 2 μl TSP 10 × PCR buffer (cat #y02026, Invitrogen, Carlsbad, CA), 0.5 μl 50 mM $MgCl_2$, 1 μl of each 10 μM detection

Table I. Primer sequences, amplicon length and LG for two markers mapped in the purine metabolism pathway of soybean.

Gene	Marker	Primer sequence 5′–3′	Amplicon size(s)	LG
Hypoxanthine phosphoribosyl transferase	Hypx #2 (F)	AGCGAGGACCAAATTTCT	650	
	Hypx #3 (R)	CTCCTTGCAGGTTTATCAAG	700	
Uricase	Uricase #9 (F)	CACAGTATAGCCTCCCTCAG	380	
	Uricase #12 (R)	TTCACAATAGGACCATCCTT		I
ASPE SNP Uricase primers	5′-LUA5(10) + TCCTGATACTTTTCTGTAAA GTAAA*A(G)*-3′			

"LUA" designates Luminex ZipCode bead sequence added to the 5′ end of the ASPE primer.

oligo, (synthesized with a unique ZipCode sequence), 0.75 U TSP DNA polymerase (cat #11448, Invitrogen), 0.1 μl 100 μM dATP, dGTP, dTTP; 0.45 μl 400 μM biotin-labelled dCTP (cat #19518-018, Invitrogen), and 11 μl of H2O. Allele-specific SNP detection PCR conditions consisted of an initial 96°C denaturation for 2 min followed by 30 cycles of denaturing (94°C for 30 s), annealing (55°C for 45 s) and extension (74°C for 45 s). A total of 80 μl of 75% ethanol was added to each 22 μl reaction, incubated in the dark for 30 min, and then centrifuged at maximum speed (4000 rpm; 3220*g*) at room temperature for 30 min. The PCR plate was immediately and carefully inverted over a paper towel and then allowed to dry upright in the dark. A total of 50 μl of TMAC buffer (3 M TMAC (Sigma T-3411), 0.1% Sarkosyl, 50 mM Tris–HCl (Sigma T-3038) and 4 mM EDTA (Gibco #15575), containing 50 beads/μl (Luminex cat #s L100-U010 (or U005)-01; Austin, TX) was added to each reaction. Reactions were then heated to 94°C for 2 min, then held at 54°C for 30 min. To each reaction, 10 μl of TMAC buffer with 0.3 μl streptavidin (Molecular Probes, Eugene OR; cat # S-866) was added and carefully mixed, followed by 15 min of 54°C. All PCR and hybridization reactions were performed on an MJ research PTC-200 thermal cycler (Biorad, San Francisco, CA). This sample was detected on a Luminex 100 instrument using the "per region" setting and Miraibio CT/GT software. Median fluorescent intensity (MFI) was used to identify specific alleles, with the lowest MFI ratio used to score a homozygous allele at 84%. A replication of the experiment was performed as above, but with a reduced concentration of 25 beads/ul of TMAC.

Map analysis

Genetic map analysis was performed using Mapmaker 3.0b (Lander et al. 1987). Both markers were tested against a 1514 Kosambi cM map containing 157 SSR molecular markers and four classical traits including height, pod color, flower color and stem termination (Shultz et al. 2007). All markers and classical traits were grouped at LOD 5.0. Both the urate oxidase and HPT markers were repeated *de novo*, scored, and then compared to verify correct allele identification.

Results and discussion

Hypoxanthine phosphoribosyltransferase

The comparison of two HPT EST sequences yielded significant similarity from bp 27–447 of Williams and bp 8–428 of Harosoy (data not shown). A single reverse primer (designated hypX-3) was designed at an identity from 442 to 422 bp of the reverse complement of the Williams sequence. Two forward primers were designed and designated as hypX-1 and hypX-2. The first was designed at the non-identical Williams sequence starting at 9 bp and extending to 27 bp. The second forward primer, hypX-2, was designed based on an identity starting at Williams bp 66, and extending for 18 bp. These three primers resulted in two primer combinations.

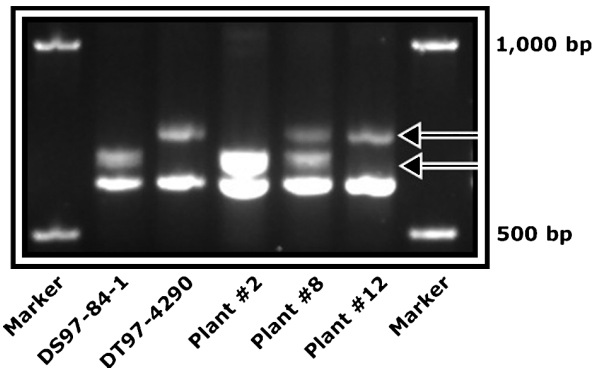

Figure 1. Gel electrophoresis of hypoxanthine phosphoribosyl transferase (HPT) marker using DT97-4290 × DS97-84-1 parents and three F_2 plants. Co-dominant alleles segregate at 700 and 650 bp, respectively (white arrows on right). Size standard (Biorad cat# 170-8361) is annotated by bp length.

The polymorphic, co-dominant HPT alleles (Figure 1) amplified by primers hypX-2 and hypX-3 (Table I) occurred at 700 (DT97-4290) and 650 bp (DS97-84-1) and mapped to linkage group (LG) "O" between Sat_478 and Sat_307. Segregation of 87 scored alleles for HPT was 19 (DT97-4290), 39 (heterozygous) and 29 (DS97-84-1), which fit the expected 1:2:1 ratio ($\chi^2 = 2.572$). Amplicons from both parents and several progeny were excised and sequenced. These sequences showed identity in comparison with soybean EST clone gi10847241 (Harosoy) annotated as similar to a HPT. Amplicons were then screened against the NCBI protein database using the tBLASTx program (6/20/2006; 4,084,561 sequences). Table II shows the top five matches for sequence from the DT97-4290 marker amplicon. The top five matches (10^{-33} to 10^{-30}) were from arabidopsis, all with HPT synonym annotation.

Urate oxidase

The single full-length cDNA (gi:170030) used to compare sequences for primer design yielded 25 primer sequences. None of the 32 resulting pairs produced visibly polymorphic amplicons. A combination of forward primer #9 (3659–3678 bp), and reverse primer #12 (4064–4045 bp) produced single, ~380 bp monomorphic amplicons for each parent, which were sequenced. These sequences were compared to each other using the BLAST 2 sequence utility and yielded an expect value of 0.0. A BLASTn search of the DT97-4290 amplicon yielded a 0.0 expect value for gi:170030 and 10^{-63} and 10^{-58}

Table II. Top translating protein blast (tBLASTx) matches for marker amplicons mapped in the purine metabolism pathway of soybean.

GI number	Description	Expect value
Hypoxanthine phosphoribosyltransferase		
11935212	*Arabidopsis thaliana* ecotype C24 hypoxanthine-guanine phosphoribosyltransferase 1 (HPRT1)	7×10^{-33}
98960956	*A. thaliana* At1g71750 (hypoxanthine phosphoribosyltransferase activity)	7×10^{-33}
21407691	*A. thaliana* clone 99557 (putative hypoxanthine ribosyl transferase)	9×10^{-33}
51968841	*A. thaliana* mRNA for putative hypoxanthine ribosyl transferase, clone: RAFL21-79-B15	1×10^{-32}
64554462	*A. thaliana* similar to the hypoxanthine-guanine phosphoribosyltransferase from *Bacillus firmus* gi1813472	1×10^{-30}
Urate reductase		
170030*	Soybean nodulin-35 (N-35) gene encoding a subunit of uricase II, complete cds	0.0
1944190	*G. max* mRNA for nodulin 35, complete cds	3×10^{-63}
1498169	*G. max* mRNA for nodulin 35	8×10^{-61}
1944192	*G. max* mRNA for nodulin 35	8×10^{-61}
170085	Soybean nodulin-35 mRNA, complete cds	2×10^{-58}

* Genbank record used to design uricase primers.

for *G. max* mRNA nodulin and urate oxidase records (Table II; 6/20/2006; 4,084,561 sequences).

Two single nucleotide differences were found within the #9 and #12 amplicon products (Table II). Polymorphic SNP amplification was recorded for the SNP located at 3685 bp of gi:170030. The SNP polymorphism mapped to LG I between Satt451 and Satt496. Segregation of 89 scored alleles for this polymorphism was 24 (DT97-4290), 45 (heterozygous) and 20 (DS97-84-1), which fit the expected 1:2:1 ratio ($\chi^2 = 0.366$). Using the manufacturers' recommended bead concentration, individual line homozygous allele (G or A) adjusted MFI ration (Figure 2), ranged from 0.91 to 0.99. Adjusted MFI ratios using the reduced-bead protocol ranged from 0.84 to 0.97 (Figure 3).

The map location of the urate oxidase SNP is verified by shared polymorphism between it and the AB002807 SSR marker. This marker amplifies a dominant polymorphism (not shown) between the two parents, with DT97-4290 lacking the genomic template. The recessive allele (lack of amplicon) segregates with the 11 homozygous DT97-4290 SNP in 21 of 22 plants. A graphical overview of these sequences in relation to gi:170030 is shown in Figure 4.

Although well covered by molecular markers, the region surrounding hypx23 on LG O is sparsely populated with QTL. Only one locus is reported, that of first flower (Yamanaka et al. 2001). The region surrounding the urate oxidase gene on LG I is well-covered by molecular markers. This area has been identified to have significant effects on protein (Sebolt et al. 2000; Chung et al. 2003), oil (Specht et al. 2001; Chung et al. 2003), plant height (Sebolt et al. 2000; Chapman et al. 2003), seed yield (Sebolt et al. 2000), leaflet area (Yamanaka et al. 2001) and pod maturity (Sebolt et al. 2000). The effect of the mapped urate oxidase and HPT genes on these traits is not known at this time.

For the urate oxidase SNP, a similar protocol was used for both the initial and confirmation population screens, with the only difference a reduction from 2500 to 1250 beads/reaction. The lowest homozygous allele adjusted MFI (84%) reflects the reduced bead concentration, and, although it appears arbitrary, Figure 3 shows that the data was discontinuous,

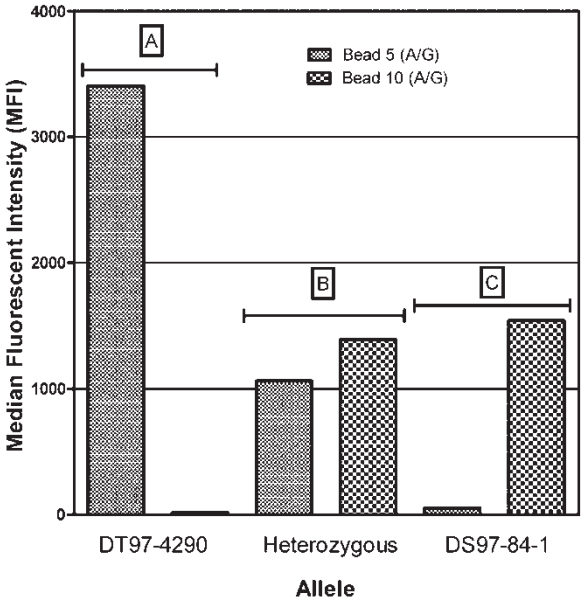

Figure 2. Detection of SNP hybridization to alternate alleles of a uricase gene in soybean. Group A is plant 20, indicative of a homozygous DT97-4290 allele. Group B is plant 47, indicative of heterozygous alleles. Group C is plant 107, indicative of a homozygous DS97-84-1 allele. Vertical bars indicate median fluourescent intensity of each allele (Bead #5 (left bar), represents DT97-4290 or "A"; Bead #10 (right bar), represents DS97-84-1 or "G").

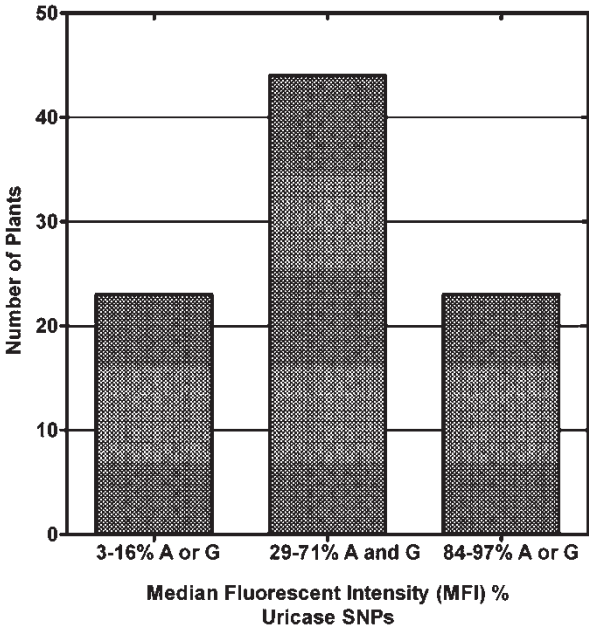

Figure 3. Discontinuous distribution of soybean Uricase SNP alleles (called by MFI ratio) in a population of 90 F_2 plants created from a cross between DT97-4290 × DS97-84-1.

clearly delimiting homozygous and heterozygous alleles. Thus, although the *ratio* was reduced by using fewer beads, the resulting data was still clearly discontinuous and called alleles exactly matched using high and low bead concentrations.

The SNP location mapped (3685 bp, gi:170030) is indicative of a difference in protein sequence (**C**UU or **U**UU), with either leucine or phenylalanine resulting, based on the change in nucleotide (alignment not shown). In contrast, the SNP located at position 3994 failed to produce reliable results, and was a silent mutation in the third bp of an alanine codon (GC**U** or GC**C**).

Variations of urate oxidase gene sequences have been reported since 1985 (Nguyen et al. 1985). Sequence analysis of the primary cDNA used for primer design (gi:170030) and the amplicons generated for SNP detection between the two parents reveals that the gi:170030 and DT97-4290 sequences are similar to the isolated N-35 nodule-specific urate oxidase gene. Shared differences with the N-35 gene allowed the identification of DS97-84-1 as similar to the isolated and described genomic clone identified by Nguyen et al. (1985). A SNP based on the gi:170030 sequence has been reported (M10594), but no information such as location of SNP detected in the sequence, primers used or map location were reported (Zhu et al. 2003). Although another marker is available in this region (AB002807), its PCR products were dominant in this population, severely limiting its usefulness for MAS.

The commercial usefulness of a molecular marker is based on its proximity to an agronomically important gene. The value of a marker is thus measured as inversely proportionate to its distance from this gene. It seems clear that once effective markers are designed that amplify polymorphism within genes, the importance of other marker types is reduced. Furthermore, a completely sequenced genome will not alleviate the need for quick, reliable testing of new varieties, a task for which gene markers are well suited. The repeat nature of the soybean genome allows the designation of each of the mapped genes as one of possibly many genes with similar sequence (Shultz et al. 2006). Genes in other locations that may have function or sequence similar to the two mapped herein have not been reported.

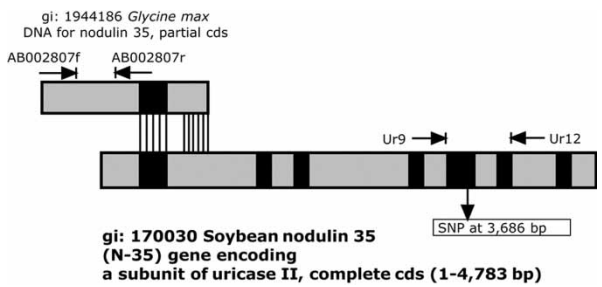

Figure 4. Graphic representation of uricase markers and their location within the target gene(s). Exonic sequence is represented by solid black fill, intronic sequence is "hashed". Similar sequences are indicated by vertical lines. Diagram is based on a full sequence alignment.

References

Atkins CA, Smith PMC. 2000. Ureide synthesis in legume nodules. Wymondham, Norfolk: Horizon Scientific Press.

Chapman A, Pantalone VR, Ustun A, Allen FL, Landau-Ellis D, Trigiano RN, Gresshoff PM. 2003. Quantitative trait loci for agronomic and seed quality traits in an F_2 and $F_{4:6}$ soybean population. Euphytica 129:387–393.

Chung J, Babka HL, Graef GL, Staswick PE, Lee DJ, Cregan PB, Shoemaker RC, Specht JE. 2003. The seed protein, oil, and yield QTL on soybean linkage group I. Crop Sci 43:1053–1067.

Kegg. 2006. Kegg database; *Glycine max* ESTs identified in purine metabolism.

King CA, Purcell LC. 2005. Inhibition of N_2 fixation in soybean is associated with elevated ureides and amino acids. Plant Physiol 137:1389–1396.

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L. 1987. Mapmaker: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181.

Nguyen T, Zelechowska M, Foster V, Bergmann H, Verma DP. 1985. Primary structure of the soybean nodulin-35 gene encoding uricase II localized in the peroxisomes of uninfected cells of nodules. Proc Natl Acad Sci USA 82:5040–5044.

Paris RL, Mengistu A, Tyler JM, Smith JR. 2006. Registration of soybean germplasm line DT97-4290 with moderate resistance to charcoal rot. Crop Sci 46:2324–2325.

Ray JD, Heatherly LG, Fritschi FB. 2006. Influence of large amounts of nitrogen on nonirrigated and irrigated soybean. Crop Sci 46:52–60.

Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. Totowa, NJ: Humana Press.

Sebolt AM, Shoemaker RC, Diers BW. 2000. Analysis of a quantitative trait locus allele from wild soybean that increases seed protein concentration in soybean. Crop Sci 40:1438–1444.

Shopinski KL, Iqbal MJ, Shultz JL, Jayakumaraman D, Lightfoot DA. 2006. Development of a pooled probe method for locating small gene families in a physical map of soybean using stress related paralogues and a BAC minimum tile path. Plant Methods 2:20.

Shultz JL, Kurunam D, Shopinski K, Iqbal MJ, Kazi S, Zobrist K, Bashir R, Yaegashi S, Lavu N, Afzal AJ, Yesudas CR, Kassem MA, Wu C, Zhang HB, Town CD, Meksem K, Lightfoot DA. 2006. The Soybean Genome Database (SoyGD): A browser for display of duplicated, polyploid, regions and sequence tagged sites on the integrated physical and genetic maps of *Glycine max*. Nucleic Acids Res 34:D758–D765.

Shultz JL, Ray JD, Smith JR, Mengistu A. 2007. A soybean mapping population specific to the early soybean production system. DNA Seq 18:104–111.

Sinclair TR, Purcell LC, Vadez V, Serraj R. 2001. Selection of soybean (*Glycine max*) lines for increased tolerance of N_2 fixation to drying soil. Agronomie 21:653–657.

Smith PM, Atkins CA. 2002. Purine biosynthesis. Big in cell division, even bigger in nitrogen assimilation. Plant Physiol 128: 793–802.

Specht JE, Chase K, Macrander M, Graef GL, Chung J, Markwell JP, Germann M, Orf JH, Lark KG. 2001. Soybean response to water: A QTL analysis of drought tolerance. Crop Sci 41: 493–509.

Takane K, Tajima S, Kouchi H. 1997. Two distinct uricase II (nodulin 35) genes are differentially expressed in soybean plants. Mol Plant Microbe Interact 10:735–741.

Todd CD, Tipton PA, Blevins DG, Piedras P, Pineda M, Polacco JC. 2006. Update on ureide degradation in legumes. J Exp Bot 57:5–12.

Vadez V, Sinclair TR. 2001. Leaf ureide degradation and N(2) fixation tolerance to water deficit in soybean. J Exp Bot 52: 153–159.

Yamanaka N, Ninomiya S, Hoshi M, Tsubokura Y, Yano M, Nagamura Y, Sasaki T, Harada K. 2001. An informative linkage map of soybean reveals QTLs for flowering time, leaflet morphology and regions of segregation distortion. DNA Res 8: 61–72.

Zhu H, Kim DJ, Baek JM, Choi HK, Ellis LC, Kuester H, McCombie WR, Peng HM, Cook DR. 2003. Syntenic relationships between *Medicago truncatula* and *Arabidopsis* reveal extensive divergence of genome organization. Plant Physiol 131: 1018–1026.